

insertions. First, interaction centers were computed from the template. Then,
 starting from the first aligned position, the lattice chain was sequentially built. At
 each step in the aligned region, the new vectors were selected so as to minimize the
 distance of the lattice chain from the equivalent template points. In the gap regions,
 the distance from the last residue of the preceding aligned fragment to the first
 residue of the next was divided to generate a set of checkpoints. The number of
 these checkpoints was equal to the number of target sequence residues that had to be
 mounted to span the gap. The checkpoints outside the entire alignment were
 generated in a random fashion. The set of all checkpoints provides the target for the
 starting lattice model. The model chain maintains the excluded volume and satisfies
 the other geometric restrictions discussed before.

Implementation of the template restraints

The template (more precisely the structural fragments of the template protein
 that correspond to the aligned residues of the probe sequence) was projected onto the
 underlying cubic lattice. The corresponding three-dimensional array, initially filled
 with zeros, was then updated to store a loose trace of the template. All elements of
 the array that were closer than $6^{1/2}$ lattice units from template residues were assigned
 the corresponding residue numbers. When a lattice point was within a distance of
 $6^{1/2}$ from two or more residues, the number of the closest residue was assigned to the
 corresponding element of the occupancy array. In the direction towards the center
 of mass of the template, the cut-off distance for creating the template "tube" was
 equal to $14^{1/2}$ instead of the $6^{1/2}$ value in the other direction. This filled in most of
 the volume occupied by the template structure. Figure 14 schematically shows such
 tubes surrounding the aligned fragments of the template chain (in solid lines). To
 illustrate the above-mentioned different width of the tube in the directions towards
 the center (versus the outside) of the template structure, the blobs forming tube were
 shifted towards the center of mass of the template. This facilitated the close packing
 of the query (target) chain that wanders within the tube.

5 As described in the previous section, the starting model was placed into the template tube. The initial alignment provided an equivalence list between the template and target residue indices. This was called "the old assignment" in contrast to the "new assignment" which was generated by the program. Both the old and the new assignments were then evaluated and updated in the following way:

- 10 a) At very beginning of the simulation process, the old assignment (the original alignment) was copied into the new assignment list. The entries of these lists identify the tube compartments and the equivalent residues of the template chain. Then, all residues for which the total number of long distance ($i-j > 4$) contacts for a three-residue fragment (with the residue of interest as a central one) was smaller than 2 become non assigned both in the old and new assignment lists. This erased those template fragments that did not interact with the rest of model protein. Thus, "non compact" fragments of the
- 15 template are ignored.
- 20 b) The new assignment was then modified when, for a steric reason (or due to local stiffness), the initial query chain residue simultaneously satisfied the following two criteria: (i) the bead of the query chain was farther away than 5 lattice units from the corresponding template residue of the original equivalence assignment ("old assignment"), (ii) the position of the query chain residue (the central point of the excluded volume cluster) coincided with a lattice point that is assigned to any other template residue. The
- 25 number read from the appropriate element (occupied by the lattice chain) of the occupancy array that corresponded to the bead coordinates became the updated entry of the new equivalence list.
- 30 c) For all residues of the starting query chain that were farther away than 9 lattice units from the equivalent (according to the old assignment) template residues, both old and new assignments were erased. These residues also became non-assigned. All allowed updates of the old assignments could only remove some entries from the equivalence list, which meant that some

part of the threading alignment was erased. The new assignments were
 5 dynamic (due to the updates described in b), and they had the character of a
 structural superposition, which was not sequential in many places.

This updated pair of assignments of the query chain residues to the template
 defined a flexible tube around the template chain. To keep the moving query chain
 in the neighborhood of the template, a set of biases was introduced. First, the model
 10 chain was kept in the broad vicinity of the original template (according to the
 updated old assignment list) by

$$E_{\text{temp},o} = \sum \delta_o(i) f_r \max \{0, (|r_i - r_{oi}| - 9)\} \quad (13)$$

where f_r was a constant (equal to $1kT$ in all simulations), r_i was the position of the
 query chain, r_{oi} was the position of the template and $\delta_o(i)$ was equal to 1 (0) when
 15 the residue i was assigned (non assigned) according the old alignment.

Then, the residues of the query chain were similarly bonded to the template
 residues in the new assignment by

$$E_{\text{temp},n} = \sum \delta_n(i) f_r \max \{0, (|r_i - r_{ni}| - R_t)\} \quad (14)$$

where r_{ni} was the position of the initial template according to the new assignment
 and $\delta_n(i)$ was equal to 1 (0) when the residue i was assigned (non-assigned)
 20 according the new assignment. The constant R_t was equal to 7 (4) when residue i
 occupied any point of the template tube (the residue was outside the tube, *i.e.*, the
 occupancy array at position r_i had value 0).

Additional restraints were the following:

$$25 \quad E_{\text{tube}} = -E^{\text{rep}} \sum \{\delta_o(i) \delta_3(i) + \delta_n(i) \delta_t(i) + \delta_n(i) \delta_c(i)\}$$

where $\delta_3(i)$ was equal to 1 when the residue i of the query chain was at a distance
 smaller than 3 lattice units from the template according to the old assignment,
 otherwise $\delta_3(i)$ equaled 0. The second component, $\delta_t(i)$, was equal to 1 (0) when the
 residue was anywhere in the template tube (is outside). $\delta_c(i)$ was equal to 1 for a
 30 "quasi-continuous" alignment on the tube, *i.e.*, when $\{al(i-1) + al(i+1)\}/2 - al(i) < 2$,
 where $al(i)$ was the value of occupancy array in the tube for residue i of the query
 chain, otherwise $\delta_c(i)$ equaled 0.